

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 4, 2002, 01:57:45 ; Search time 24.7555 Seconds  
(without alignments)  
1429.075 Million cell updates/sec

Title: US-09-805-550-4

Perfect score: 1910  
Sequence: 1 MKLTVKTLKGFHFEIRVQPN.....CDRNEELANYLLHAGGED 368

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1153.5	60.4	379	2 T14337	RAD23 protein, iso
2	1043.5	54.6	367	2 F96827	protein F20B17.8 f
3	954	49.9	382	2 T14336	RAD23 protein, iso
4	887.5	46.5	392	2 T04150	RAD23 protein, iso
5	661.5	34.6	246	2 G86296	T24D18.27 protein
6	640.5	33.5	385	2 JC7783	RAD 23B protein -
7	625.5	32.7	409	2 S44346	RAD23 protein homo
8	603.5	31.6	363	2 S44443	RAD23 protein homo
9	561.5	29.4	368	2 T40115	uv excision repair
10	424	22.2	113	2 H86296	F309.1 protein - A
11	405	21.2	372	2 T27774	hypothetical prote
12	403	21.1	398	2 S50507	excision repair pr
13	204.5	10.7	142	2 T51479	hypothetical prote
14	201	10.5	551	2 C84549	probable ubiquitin
15	195.5	10.2	536	2 B84549	probable ubiquitin
16	163.5	8.6	354	2 T38404	yeast dsk2 homolog
17	163.5	8.6	502	2 T18562	hypothetical prote
18	159	8.3	373	2 S54583	ubiquitin-like pro
19	143	7.5	324	2 T02358	ubiquitin homolog
20	140	7.3	701	2 T06674	ubiquitin-like fus
21	138	7.2	313	2 T01728	ubiquitin precursor
22	137.5	7.2	292	2 T18584	F15C11.2 - Caenorh
23	136	7.1	1135	2 T30561	Scytin protein - A
24	135	7.1	356	1 U0079C	polyubiquitin / ri
25	132.5	6.9	990	2 T14756	hypothetical prote
26	130.5	6.8	76	2 S31653	ubiquitin - giardi
27	127.5	6.7	80	2 T28305	ORF MSV144 probabl
28	126.5	6.6	77	2 S29238	ubiquitin - Entamo
29	126.5	6.6	156	1 U0H0R7	ubiquitin / riboso

30	126.5	6.6	156	1 I52328	ubiquitin / riboso
31	126.5	6.6	157	2 I50438	ubiquitin polyprot
32	126.5	6.6	167	2 I51568	ubiquitin - Af
33	126.5	6.6	222	1 U0H7	ubiquitin precursor
34	126.5	6.6	229	1 U0H7	ubiquitin precursor
35	126.5	6.6	229	1 S13928	ubiquitin precursor
36	126.5	6.6	231	1 A31560	polyubiquitin - It
37	126.5	6.6	245	2 A29584	ubiquitin - bo
38	126.5	6.6	255	2 A29584	ubiquitin precursor
39	126.5	6.6	305	1 S29853	polyubiquitin 4 -
40	126.5	6.6	305	1 I50437	polyubiquitin 4 -
41	126.5	6.6	305	1 S12583	polyubiquitin 4 -
42	126.5	6.6	381	2 S21083	polyubiquitin 5 -
43	126.5	6.6	685	1 U0H0C	polyubiquitin 9 -
44	126.5	6.6	810	2 S45359	polyubiquitin 10 -
45	126	6.6	155	1 U0W07A	ubiquitin / riboso

## ALIGNMENTS

RESULT 1  
T14337  
RAD23 protein, isoform II - carrot  
C:Species: Daucus carota (carrot)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: T14337  
R:Sturm, A.; Reinhard, S.  
Plant J. 13, 815-821, 1998  
A:Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.  
A:Reference number: Z17989; MUID:98345997; PMID:9681019  
A:Accession: T14337  
A:Status: preliminary; translated from GH/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-379 <SPU>  
A:Cross-references: EMBL:Y12014; NID:G1914684; PIDN:CAA72742.1; PID:G1914685  
A:Experimental source: subspecies Queen Anne's Lace, isolate W001C  
A:Genetics:  
A:Gene: RAD23-2  
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology.

Query Match	Score	DB 2;	Length
Best Local Similarity	60.48;	Score 1153.5;	DB 2; Length 379;
Matches 240; Conservative 45; Mismatches 81; Indels 15; Gaps 9;			
QY 1	1	MKLTVKTLKGFHFEIRVQPNDTMAVKNIEIQKSYPMGQQLIFNGVKLDESTLE 60	
DB 1	1	MKLTVKTLKGFHFEIRVQPNDTMAVKNIEIQKSYPMGQQLIFNGVKLDESTLE 60	
QY 61	61	ENKYNEDGFLVYMLSKGTSSTSSQSHSNTPATROAP---PLEAPQAPPPVAPIT 117	
DB 61	61	ESKISSEGFLVYMLSKGTSSTSSQSHSNTPATROAP---PLEAPQAPPPVAPIT 120	
QY 118	118	TSQPEG--LPAQAP-NTHDNAASNLGSRVDTIINDLEMGGSSWCKDYQALRAAYN 174	
DB 121	121	T-VEEAPLSPAFAPSDTYGEAASNVVAGSNLEQIIOHIMDGGCMPTNNVSRALRAAYN 179	
QY 175	175	NPEAAVEYVSGIVTAEIANPDI---GQGANITDRA--PRGEGAGLS-GIPNTAPDLFP 228	
DB 180	180	NPEAAVDTLYSGIPEMEAAPVSHFQGDQINAAANNISDNGVGAAPGAGANSPLMNF 239	
QY 229	229	QGASNAGGAGGGLDFLRNNPQFAVENVHTNPQILQPMVLVLSKONPQILILEENH 288	
DB 240	240	QETLSGVYAGAGLSLEFLRNPNQFQTLRSWQVRNPQILQPMVLVLSKONPQILQIOGHN 299	
QY 289	289	DEFTQLNEPPEGGEFLDQPEEDENPHASVTPPEDEAIGRLSESGFDRARYEAPLA 348	
DB 300	300	EEFQQLNEPPEASEGDFDQPEED-VQETVTAADQEAERLEANGFDGLVTEAFLA 358	
QY 349	349	CDRNEELANYLLHAGG-ED 368	
DB 359	359	CDRNEELANYLLHAGGDFED 379	

## RESULT 2

protein F20B17.8 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002

C:Accession: F96827

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpspiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F96827

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-367 <STO>

A:Cross-references: GB:AE005173; NID:97715605; PIDN:AAF68123.1; GSPDB:GND00141

C:Genetics:

A:Gene: F20B17.8

A:Map position: 1

C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 54.6%; Score 1043.5; DB 2; Length 367;

Best Local Similarity 59.3%; Pred. No. 1.2e-62;

Matches 232; Conservative 39; Mismatches 73; Indels 47; Gaps 11;

1 MKLYVTKLGFHEIRVOQNDITMAVKNIEIOGDSYPMGQOILLFNGKYLKDESTLE 60

1 MKLYVTKLGFHEIRVOQNDITMAVKNIEIOGDSYPMGQOILLFNGKYLKDESTLE 60

61 ENKVNEDGFLVYMLSKGKTSSTSSQ-----HSNTPATRQAPLEAQAQAPPV 113

61 ENKVNEDGFLVYMLSKGKTSSTSSQ-----HSNTPATRQAPLEAQAQAPPV 113

61 ENKVNEDGFLVYMLSKGKTSSTSSQ-----HSNTPATRQAPLEAQAQAPPV 113

114 APTT-----SOEGLPAPQAPNTHDNANSLSGRNVDTIINQLEMGGSGMDKRV 165

112 NFVFTCSCTFYSCSRTTG-----TDYGAQASTLVSGSLQEMVOQIMEMGGSGMDKRV 166

166 GQALAAVNNPRAVEYLYSGIPVTAETV-----IGQGANTDRATGAGLSGIRN 220

167 TALRAATNPPRAVDYLYSGIPVTAETV-----IGQGANTDRATGAGLSGIRN 219

221 TAPLDFPQAGSNAGGAGGPPDLFLRNPPQAVREMYHTNPQILQPMIVELSKONPQI 280

220 SSPDLDFPQETVAAGSGDILGTFELRNNDQVAIL--TISAFSLNCEPRMLQELGKQNPQ 277

281 LALIEBNHDEFQILNEPREG--GGGDFLDQPEDEMPRAISVTPPEQAIQGLSMD 338

278 LALIEBNHDEFQILNEPREG--GGGDFLDQPEDEMPRAISVTPPEQAIQGLSMD 336

339 RARVIEAFLACDRNEELANYLEHAGE-ED 368

337 RALVIEAFLACDRNEELANYLEHAGE-ED 367

RESULT 3

714336

RAD23 protein, isoform I - carrot

C:Species: Daucus carota (carrot)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000

C:Accession: T14336

R:Sturm, A.; Leibhard, S.

A:Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.

A:Reference number: Z17989; MUID:96345997; PMID:9681019

A:Accession: T14336

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-382 <STO>

A:Cross-references: EMBL:Y12013; NID:91914682; PIDN:CAA72741.1; PID:91914683

A:Experimental source: subspecies Queen Anne's lace, isolate W001C

C:Genetics:

A:Gene: RAD23-1

C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 49.9%; Score 954; DB 2; Length 382;

Best Local Similarity 52.4%; Pred. No. 1.3e-56;

Matches 205; Conservative 52; Mismatches 98; Indels 36; Gaps 9;

1 MKLYVTKLGFHEIRVOQNDITMAVKNIEIOGDSYPMGQOILLFNGKYLKDESTLE 60

1 MKLYVTKLGFHEIRVOQNDITMAVKNIEIOGDSYPMGQOILLFNGKYLKDESTLE 60

61 ENKVNEDGFLVYMLSKGKTSSTSSQHSNTPATRQAPLEAQAQAPPV-OPVPA- 114

61 ENKVNEDGFLVYMLSKGKTSSTSSQHSNTPATRQAPLEAQAQAPPV-OPVPA- 114

61 ENKVNEDGFLVYMLSKGKTSSTSSQHSNTPATRQAPLEAQAQAPPV-OPVPA- 114

115 -PITTSQPGGLPAQAP-----NTHDNANSLSGRNVDTIINQLEMGGSGMDK 163

117 LPVPAPAPAPATAPIPSAAGSEANVYDSASILVAGSINLEGAIQILLDMGGGTWDRD 176

164 KYORALRAAVNNPRAVEYLYSGIPVTAETV-----GGGANTDRATG--EAGIS 216

177 TVIRIVRAAVNNPRAVEYLYSGIPVTAETV-----GGGANTDRATG--EAGIS 235

217 GIPNTAPLDFPQAGSNAG--GGAGGPPDLFLRNPPQAVREMYHTNPQILQPMIVELSK 275

236 AGPNAPNAPLDFPQAGSNAG--GGAGGPPDLFLRNPPQAVREMYHTNPQILQPMIVELSK 295

276 QNPQILRLIEBNHDEFQILNEPREG--GGGDFLDQPEDEMPRAISVTPPEQAIQGLSMD 335

296 QNPQILRLIEBNHDEFQILNEPREG--GGGDFLDQPEDEMPRAISVTPPEQAIQGLSMD 349

336 GFDRAVIEAFLACDRNEELANYLEHAGE 366

350 GFDRELIVLEVFRCNKEELANYLEHAGE 380

RESULT 4

704150

RAD23 protein homolog - rice

C:Species: Oryza sativa (rice)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000

C:Accession: T04150

R:Schultz, T.F.; Quatrano, R.S.

A:Title: Characterization and expression of a rice RAD23 gene.

A:Reference number: Z08695; MUID:97369378; PMID:9225866

A:Accession: T04150

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-392 <SCH>

A:Cross-references: EMBL:U63530; NID:91488296; PIDN:AA65841.1; PID:91488297

A:Experimental source: cv. Nipponbare

C:Genetics:

A:Gene: RAD23

C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 46.5%; Score 887.5; DB 2; Length 392;

Best Local Similarity 50.9%; Pred. No. 3.8e-52;

Matches 199; Conservative 52; Mismatches 109; Indels 31; Gaps 10;

1 MKLYVTKLGFHEIRVOQNDITMAVKNIEIOGDSYPMGQOILLFNGKYLKDESTLE 60

1 MKLYVTKLGFHEIRVOQNDITMAVKNIEIOGDSYPMGQOILLFNGKYLKDESTLE 60

61 ENKVNEDGFLVYMLSKGKTSSTSSQHSNTPATRQAPLEAQAQAP--OPVAPITTT 118

61 ENKVNEDGFLVYMLSKGKTSSTSSQHSNTPATRQAPLEAQAQAP--OPVAPITTT 118

61 ENKVNEDGFLVYMLSKGKTSSTSSQHSNTPATRQAPLEAQAQAP--OPVAPITTT 116

QY 119 -----SOEGCLPAQAP-----NTHDNASNLISGRNVDTITINQIMEMGGSGWDK 163  
 Db 117 VPTVSAPTPTAFASPAVASSEADNYSNVLAVASNLVETIQSIILEMGGIMDRD 176  
 QY 164 KVORALRAAYNPERAVEYLYSGIPVTAETAVPIGQ-OGANTTDRA----PTGAGLSGI 218  
 Db 177 IYVLAISAANNRERAVEYLYSGIPVTAETAVPIGQ-OGANTTDRA----PTGAGLSGI 236  
 QY 219 PNTAPLDLPFGASNAGGAGG-GPLDPLRNPOFAVREMTNPQILQPMLEYSKON 277  
 Db 237 PNASPLDLPFOALPNASTDAGLGNLDALRNNAQFRTLSIVQNPQILQPLQELKON 296  
 QY 278 PQILRIEENHDEFLQILNPEFG-GEQDPLDQ-PEEDEMPHALSYVPEDEAGRLISM 335  
 Db 297 PQILQIQEQAFLHILNPEAGDEENLDQPE--AMFQITAVPEDEALIRLEPM 354  
 QY 336 GFDRARVTEAFACDRNEELANYLEHAGE 366  
 Db 355 GFDRAVLVDVFFACNKDEQLAANTLLDHME 385

## RESULT 5

G86296  
 T24D18.27 protein - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 27-Nov-2001  
 C:Accession: G86296  
 R:Theologas, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Com. L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hutzar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maitl, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: G86296  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-246 <STO>  
 A:Cross-references: GB:AEO05172; NID:g6587822; PIDN:AAF16513.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 34.6%; Score 661.5; DB 2; Length 246;  
 Best Local Similarity 56.7%; Pred. No. 2.9e-37;  
 Matches 143; Conservative 29; Mismatches 63; Indels 17; Gaps 4;

QY 1 MKLVKTLKGFHEIRVOPNDTIAVKNKIEIGKDSYPMGQQLIFNGKVLKDESTLE 60  
 Db 1 MKLVKTLKGFHEIRVOPNDTIAVKNKIEIGKDSYPMGQQLIFNGKVLKDESTLE 60  
 QY 61 ENKNEDEGLVYMLSKGTSGSTSSQSNTPATROAPPLLEAPQAPPPVAPITTSQ 120  
 Db 61 ENKNEDEGLVYMLSKGTSGSTSSQSNTPATROAPPLLEAPQAPPPVAPITTSQ 120  
 QY 121 PEGPAQAPNTHDAASNLISGRNVDTITINQIMEMGGSGWDKQVORALRAAYNPERAV 180  
 Db 109 VQEPQASDITVGAASLTVSGSSIEQVQOIMEMGGSGWDKQVORALRAAYNPERAV 168  
 QY 181 EYLYSGIPVTAETAVPIGQ-OGANTTDRA-PTGAGLSGIPTAPLDLPFGASNAGGAG 239  
 Db 169 EYLYSGIPVTAETAVPIGQ-OGANTTDRA-PTGAGLSGIPTAPLDLPFGASNAGGAG 224  
 QY 240 GGPDLFARNPQ 251  
 Db 225 LGTLEFLRGNDQ 236

RESULT 6  
 JC7783

RAD 23B protein - channel catfish  
 C:Species: Ictalurus punctatus (channel catfish)  
 C>Date: 02-Apr-2002 #sequence\_revision 02-Apr-2002 #text\_change 02-Apr-2002  
 C:Accession: JC7783  
 R:Li, Z.; Li, P.; Kocabas, A.; Karsi, A.; Ju, Z.  
 Biochem. Biophys. Res. Commun. 289, 317-324, 2001  
 A:Title: Microsatellite-containing genes from the channel catfish brain: Evidence of  
 A:Reference number: JC7783  
 A:Contents: Brain  
 A:Accession: JC7783  
 A:Molecule type: mRNA  
 A:Residues: 1-385 <LNU>  
 C:Comment: This protein with a polypeptide tract, has importance in the nucleotide  
 C:Genetics:  
 A:Gene: rad23b  
 A:Introns: 76/73

Query Match 33.5%; Score 640.5; DB 2; Length 385;  
 Best Local Similarity 35.1%; Pred. No. 1.3e-35;  
 Matches 142; Conservative 88; Mismatches 118; Indels 57; Gaps 9;

QY 1 MKLVKTLKGFHEIRVOPNDTIAVKNKIEIGKDSYPMGQQLIFNGKVLKDESTLE 60  
 Db 1 MQTLKTLQOQTFKIDDAETVAKLEKIEKESFPYAGCKLYAGKILNDTALK 60  
 QY 61 ENKNEDEGLVYMLSKGTSGSTSSQSNTPATROAPPLLEAPQAPPPVAPITTSQ 114  
 Db 61 EYKIDENKVFVYAKKRAATMAOSSSTTAASSSTSTTPVPPVAAATTTT 120  
 QY 115 PITTSQ-----PEGCLPAQA-----PETHDNASNLISGRNVDTITINQIM 153  
 Db 121 TTTTSTSESSVIEEKAEEKPSTSPASGSLTNVINFERTSALVTGGSYEMVVEIM 180  
 QY 154 EMGGSGWDKQVORALRAAYNPERAVEYLYSGIPVTAETAVPIGQAGANTTD 206  
 Db 181 LMG---YERREYVAAALASFPNPDRAVEYLYTGIPVSEGVNAGSDPAAPVAGAPATTT- 236  
 QY 207 RAPYGEAGLSGIPNTAPLDLPFGASNAGGAGGAPLDPLRNPOFAVREMTNPQIL 266  
 Db 237 -----GLSSPSTTP-----TQTAGSGGANPLEFLRNQOFLQMRQIIOQNSLS 281  
 QY 267 QPMLEYSKONPQILRIEENHDEFLQILNPEFG-GEQDPLDQ-PEEDEMPHALISVTP 323  
 Db 282 PALLQOIGRENPPOLLQISSHQEQIQLNPEPAEAGGGGVSSEAGGGMNY-IGVTP 340  
 QY 324 EEOGAIGLESMPDRARVTEAFACDRNEELANYLEHAGEED 368  
 Db 341 QEKAIERIKALGPEGLVIAQYFACRKNEMLANPILQONFDD 385

## RESULT 7

S44346  
 RAD23 protein homolog - human  
 N:Alternate names: p58/HR23B; XP-C repair complementing protein  
 C:Species: Homo sapiens (man)  
 C>Date: 25-Dec-1994 #sequence\_revision 01-Sep-1995 #text\_change 24-Sep-1999  
 C:Accession: S44346  
 R:Mesutani, C.; Sugawara, K.; Yanagisawa, J.; Sonoyama, T.; Ui, M.; Enomoto, T.; Taki,  
 EMBO J. 13, 1831-1843, 1994  
 A:Title: Purification and cloning of a nucleotide excision repair complex involving t  
 A:Reference number: S44345; MUID:94222030; PMID:8168482  
 A:Accession: S44346  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-409 <MAS>  
 A:Cross-references: EMBL:ID1090; NID:g498147; PIDN:BA004652.1; PID:dl005181; PID:g498  
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology  
 F:1-80/Domain: ubiquitin homology <UBH>

Query Match 32.7%; Score 625.5; DB 2; Length 409;  
 Best Local Similarity 34.8%; Pred. No. 1.5e-34;  
 Matches 147; Conservative 83; Mismatches 125; Indels 67; Gaps 11;



```
QY 171 AATNPRANEVLYSGIP-----VTAETAVPIGGAGANTTDRAPTGEAGLSGIPNTAPL 224
Db 170 AATNPRANEVLYSGIP-----VTAETAVPIGGAGANTTDRAPTGEAGLSGIPNTAPL 224
QY 225 DLEPQGA-----SNAGGAGGGLDFLRNPNFOFARVMTHTNPQILPMLVELSKON 277
Db 222 NLEPQALSNENQOEPSNTVGDPLGLFSLRIPQFQOLROIVQONPQMLTTLTQIQGD 281
QY 278 PQLRLTEENHDFLQILNPEFEGEGDFLDOPEDEMPHA---ISVPEEOEAIGRLS 334
Db 282 PALAQATQNPFAFLQILNPEFEGEGDFLDOPEDEMPHA---ISVPEEOEAIGRLS 334
QY 335 MGFDRARVIAFLACDRNEELANYLEHAGEED 368
Db 332 LGFDRNIVIAQVLAQCKNEELANYLEHAGEED 365

RESULT 10
H86296
F309.1 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 24-Nov-2001
C:Accession: H86296
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.M.; Chung, M.K.; Conn, E.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.: Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MID:21016719; PMID:11130712
A:Accession: H86296
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <STO>
A:Cross-references: GB:AEO05172; NID:9496345; PIDN:AAD34676.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 22.2%; Score 424; DB 2; Length 113;
Best Local Similarity 73.7%; Pred. No. 8.3e-22;
Matches 84; Conservative 16; Mismatches 10; Indels 4; Gaps 3;

QY 258 MWHNPQILPMLVELSKONPQILRLTEENHDFLQILNPEFEGEG--DFLDQPEDEM 315
Db 1 MWHNPQILPMLVELSKONPQILRLTEENHDFLQILNPEFEGEG--DFLDQPEDEM 315
QY 316 PHAIVTPEEOEAIGRLSNGFDRARVIAFLACDRNEELANYLEHAGE-ED 368
Db 60 PHAIVTPEEOEAIGRLSNGFDRARVIAFLACDRNEELANYLEHAGE-ED 368

RESULT 11
T27774
hypothetical protein ZK20.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T27774
R:Gajasty, S.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z20417
A:Accession: T27774
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-372 <MIL>
A:Cross-references: EMBL:Z69904; PIDN:CAA93780.1; GSPDB:GN00020; CESP:ZK20.3
A:Experimental source: clone ZK20
C:Genetics:
A:Gene: CESP:ZK20.3
A:Map position: 2

A:introns: 49/2; 73/3; 101/3; 369/3
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
Query Match 21.2%; Score 405; DB 2; Length 372;
Best Local Similarity 30.0%; Pred. No. 7.8e-20;
Matches 111; Conservative 65; Mismatches 142; Indels 52; Gaps 8;

QY 1 MKLTVKTLKGTHTFIRVQPDITMAVKNIEIOGKSPWGOQLTFNGKVLKDESTLE 60
Db 52 LSVTFRTLVQVNFLEINEDQITAEVKAIVASEKG--DDVAPELQKILYNGKIILDDSVK 110
QY 61 ENKYNEDGFLVWLSKGTSGTSSQHSNTPATQAPPLAPQAPPPVAPITTSQ 120
Db 111 EYVGDSSKFFVYVMSKRRVTEVAPSSVTAAPAEVPAAPASNP--APADAAPAEAAA- 167
QY 121 PEGIPAPAPNTHDNAANLLSGRNVDTITQNLNEMGGSGMDKQORALAAVNPRAV 160
Db 168 ----PAPAEALTDQENNVLA-----ITMGTDREQTIALRAAENPDRAY 210
QY 181 EYLYSGIPVTAETAVPIGGAGANTTDRAPTGEAGLSGIPNTAPDLFPQASNAGGAGC 240
Db 211 EFLNLGHPDA-----ADQEP-----DLGPEQNIIDVDEGN 242
QY 241 GPDLFLNNNQFQAVREMTHTNPQILPMLVELSKONPQILRLTEENHDFLQILNPEFE 300
Db 243 DDLMLNMLNPOLAEIRALIQONPEMLAAVLAQVLAANPRVOTTONNQAFMDLLNGAQ 302
QY 301 GGEEDFLDQPEED-EMPHATVPEEOEAIGRLS--MGFDRARVIAFLACDRNEELAA 357
Db 303 GAGCAAGNAPERNPRRHVHLSPSEAAIERIKATVYVNAPEAVVATYACDRNEEAAI 362
QY 358 NYLLEHAGEE 367
Db 363 NFIRSNIDEE 372

RESULT 12
S50507
excision repair protein RAD23 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YEL037C
C:Species: Saccharomyces cerevisiae
C>Date: 28-Jan-1995 #sequence_revision 12-May-1995 #text_change 24-Sep-1999
C:Accession: S50507; S30845; A54599; S38547
R:Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 8199, 8334, and 9871.
A:Reference number: S50491
A:Accession: S50507
A:Molecule type: DNA
A:Residues: 1-398 <DIE>
A:Cross-references: EMBL:U18779; NID:9603625; PIDN:AAB65005.1; PID:9603642; MIPS:YEL0
R:Wattkins, J.F.; Sung, P.; Prakash, L.; Prakash, S.
Mol. Cell. Biol. 13, 7757-7765, 1993
A:Title: The Saccharomyces cerevisiae DNA repair gene RAD23 encodes a nuclear protein
A:Reference number: A54599; MUID:94067136; PMID:8246991
A:Accession: A54599
A:Molecule type: DNA
A:Residues: 1-398 <NAT>
A:Cross-references: GB:U25428; NID:9409246; PIDN:AAA16070.1; PID:9409247
R:Wattkins, J.F.; Sherman, F.
J. Mol. Biol. 233, 372-388, 1993
A:Title: The gene clusters ARC and COR on chromosomes 5 and 10, respectively, of Sacc
A:Reference number: S38543; MUID:94016558; PMID:8411151
A:Accession: S38543
A:Molecule type: DNA
A:Residues: 1-276, 278-398 <MEL>
```



B84549  
 Probable ubiquitin-like protein [Imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: B84549  
 R;lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: B84549  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-536 <STO>  
 A:Cross-references: GB:AE002093; NID:g4584342; PIDN:AAD25137.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g17190  
 A:Map position: 2

Query Match 10.2%; Score 195.5; DB 2; Length 536;  
 Best Local Similarity 20.3%; Pred. No. 1.4e-05;  
 Matches 110; Conservative 60; Mismatches 170; Indels 201; Gaps 19;

```

QY 3 LTVKTLGTHPEIRVQPNQITMAVKNIIEIQKDSIPWGOQLIFNGKVLKDSSTLEEN 62
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 20 VNVKSCNGTKFSTYTSIDSTVESFK--ELIAQNSDVPANQRLIKGRILKDDOTLLSY 76
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 63 KVNEDGFLVYVLSKSGKTSSTGTSQHSNTPATROA----- 99
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 77 GLQADH--YVAMVGRGFPVSSPSAPANAQNTAPQAVGSNDSSNLGGESLFPGLGFNP 134
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 100 -----PLEAPQAPPPVAPITTSOPEGLPA--QAPYTHDMAASNLG 142
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 135 LGGGNMAGLEFSGSLPRLDQAOQ--QLAQNPMIREMNTPAIQMLNMPFMSMINN 192
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 143 RNVDITINQLEMGGSGNDKQVQALRAVYN-----NPERA---VELYSGIPVT 190
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 193 PQRELVDNRNDELGVNDPSILRQLEAARNPELMRMNRTDRAMSNIESMPEGFNWL 252
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 191 AEI-----AVPTGGGAGNT-----TDRAPTEGA 213
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 253 RRMENVOEPLMATTSNAGNNTSSNPFALLGNQGVTTQGSPTSNNISAPNAETGTP 312
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 214 GLSGIPN-----TAPLDLPQASNAG-----GGAG-----G 241
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 313 NANPLPNWGAITAGOTTP-----GRTNAGLGGIGLGGIGLGMGADSPGATPDAS 366
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 242 PLDFLRNNPQFAVRENVHTNPQLIQMLVELSKNPQILRLIEBHDEFLQLLNEPREG 301
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 367 QLSQILQNPAMQSVLSNPQYN---QLMSLNPQLRSMIDWN--POLREMGNP--- 418
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 302 GEGDFLOPDEDEMPHAI----- 319
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 419 --DFLRQFSSPEEMQMSLQOSLFSONRNTAGODPTQGAATANNGLDLMMNFGSL 475
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 320 -----SYTPEQEA--IGRLSEMGF--DRARYIEAFIACDRNEELANYLLERHAG 365
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 476 GAGGLSGTNGPNVPEERFAIQLOQLQEMGFYDRAENIRALLATNGNNAVERLLGSIG 535
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 366 E 366
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 536 Q 536
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

Search completed: November 4, 2002, 02:03:51  
 Job time : 26.7555 secs

